Report on the First Working Group Meeting of the “AG Marketing”

Friederike Paetz, Andreas Geyer-Schulz and Winfried J. Steiner

Abstract This contribution reports on the first meeting of the new formed working group “Data Analysis and Classification in Marketing (AG Marketing)” of the data science society (GfKl) held at the KIT, Karlsruhe, November 14th – 15th, 2019. The abstracts of the presentations given reflect the ongoing trend to exploit a large variety of digital data sources for marketing purposes and the need for advanced and innovative analysis methods.

Friederike Paetz
Technical University Clausthal, Julius-Albert-Str. 6, D-38678 Clausthal-Zellerfeld,
✉ friederike.paetz@tu-clausthal.de

Andreas Geyer-Schulz
Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, D-76131 Karlsruhe Germany
✉ andreas.geyer-schulz@kit.edu

Winfried J. Steiner
Technical University Clausthal, Julius-Albert-Str. 6, D-38678 Clausthal-Zellerfeld,
✉ winfried.steiner@tu-clausthal.de

Archives of Data Science, Series A
KIT Scientific Publishing
Vol. 7, No. 1, 2020
DOI: 10.5445/IR/1000119609
ISSN 2510-0564

Introduction

We are happy to announce the formation of the new working group "AG Marketing" under the umbrella of the Data Science Society. The AG Marketing focuses on quantitative marketing research and bundles the competencies of academics and practitioners from the marketing sector. Since November 2019 the working group is led by PD Dr. Friederike Paetz, Clausthal University of Technology, and by Prof. Dr. Daniel Guhl, Humboldt University Berlin, as deputy head. The first meeting of AG Marketing took place from 14th to 15th of November 2019 at the KIT Karlsruhe Institute of Technology in Karlsruhe, Germany.

Nowadays, quantitative marketing research is of high importance for both marketing academics and marketing practitioners. The knowledge of data analytics and classification patterns in marketing contributes to successful marketing decisions that are based on sophisticated tools for data science. In particular, the development of quantitative marketing models and advanced quantitative methods for data analysis in a marketing context is essential for extracting and gaining marketing-relevant information from a wide variety of data sources. One of the challenges for marketing researchers will be the integration of sophisticated pre-processing algorithms and pre-trained knowledge-bases from machine learning and artificial intelligence for an improved understanding of natural language, video and audio-sources.

The application of advanced techniques is promising for practical marketing decisions in varying areas like direct marketing campaigns (Baier, D. & Stöcker, B.), sales management (Rausch, T., Albrecht, T. & Baier, D.), sales interaction (Pauser, S. & Wagner, U.), dynamic pricing (Ascherleben, P. & Steiner, W.J.), digital advertising (Furtado, F.S., Reutterer, T. & Schröder, N.), user comment analysis (Hartmann, J., Schwenzow, J. & Schikowsky, A.), consumer behavior (Yegoryan, N., Guhl, D., Paetz, F. & Klapper, D.), the discovery of consumer preferences (Laghaie, A. & Otter, T.), or the characterization of target consumers (Paetz, F.). In addition, the investigation of conceptual issues for the application of data analyses such as the handling of missing values in data (Grimm, M.S. & Wagner, R.) or the requirements for specific model applications (Simon, L.) contributes to a deeper understanding of quantitative marketing research techniques.

The first working group meeting of AG Marketing provided a wide overview of quantitative marketing research techniques in both concept-oriented as well as application-oriented studies. The abstracts of the contributions that were presented at the first working group meeting of the AG Marketing are provided in the following.

1 Maximizing Return on Investment from Direct Marketing Campaigns: A New Uplift Modeling Approach for Online Shops

Daniel Baier, Björn Stöcker

In order to improve return on investment from direct marketing campaigns, usually, a (small) sample of customers is testwise contacted and their positive reactions (e.g. bought advertised products in a predefined time slot) and negative reactions (e.g. did not buy) are used to develop a predictive response model (based e.g. on past information and buying behavior) for all customers. Then, the latter is used to select customers for the direct marketing campaign according to the highest positive response predictions among all customers. However, this classical approach has two major shortcomings: First, the response model also selects customers who would positively respond regardless of the campaign (waste of money). Second, the response model only reflects a binary outcome (bought or did not buy), not a continuous outcome (probability/propensity of buying, sales or profit). Both shortcomings restrict the usefulness of the approach when maximizing the return on investment from the direct marketing campaign. In this paper we propose a new approach that is able to overcome the discussed problems. The new approach connects findings from the field of uplift modeling (see, e.g., Radcliffe and Surry (1999); Surry and Radcliffe (2011); Kane et al (2014)) with findings from the field of sample selection (see, e.g., Heckman (1979)). Using the well-known Hillstrom data set and an own actual online shop direct marketing campaign data set (with data from >270k customers) as examples, we show that the new approach is well suited to correctly select the "right" customers as targets and to improve return on investment from direct marketing campaigns.

2 Forecasting Sub-Daily Call Center Arrivals: Investigating the Joint Impact of Data Disaggregation and Model Selection on Accuracy

Theresa Rausch, Tobias Albrecht, Daniel Baier

Customers' perception of high service quality contributes to customers' loyalty and, therefore, drives a company's success and survival within their competitive environment. Drawing on marketing literature, perceived service quality is determined by interaction quality and outcome quality. The latter comprises – among others – customers' waiting times. Thus, call center managers are expected to provide high service quality by decreasing waiting times and simultaneously to keep operating costs at a minimum by deploying an appropriate number of agents. Hence, this paper conducts a model comparison to predict call arrivals with multiple seasonality. We compare traditional and barely investigated time series models (i.e. ARIMA, Random Walk, TBATS, Innovation State Space, Dynamic Harmonic Regression), regression models (i.e. Generalized Linear Models, Zero Inflated Models), and a machine learning approach (i.e. Random Forest). Additionally, we consider a new data processing related approach to enhance forecast accuracy: We investigate whether aggregating sub-daily data to daily values and in turn, disaggregating daily predictions according to averaged call distribution per interval yields more accurate forecasts than predictions of sub-daily data. We analyze call arrivals recorded at a German online retailer's call center comprising 174.5 weeks of half-hourly data. We calculate forecast accuracy using cross validation in combination with a rolling forecast origin for 52 weeks. Our findings indicate that a Dynamic Harmonic Regression model has substantial predictive potential for practical use. Random Forest yields comparable results and outperforms traditional approaches. Moreover, we prove that time series models without explanatory variables perform more accurate on ordinary weekdays whereas machine learning and regression models with explanatory variables are more suitable to capture the course of special days, e.g., holidays. For the majority of the models, disaggregated daily predictions generate higher accuracy than predictions of sub-daily data.

3 How accurate are customers’ initial impressions? Using continuous-response measurement to assess thin slices of sales behaviors
Sandra Pauser, Udo Wagner

A good first impression is crucial for the success of a sales interaction. Prior research demonstrates that individuals are able to make accurate predictions about one’s personality, skills, traits, or competencies from brief observations, so-called thin slices. Specifically, studies point on the importance of nonverbal cues (i.e., facial expressions, gestures) in the formation of initial impressions. However, these behaviors are perceived mainly unconsciously, which makes measurement a difficult task. Moreover, existing research is dominated by post-exposure measures and neglects customers’ processing of impressions over time. This research tackles the problems outlined above and introduces continuous measurement of initial impressions in a sales context by a variety of different data sources. We provide novel insights by applying high-precision coding of nonverbal behaviors in 22 videotaped sales presentations (elevator pitches) by making use of the body action and posture coding procedure (BAP), which allows the analysis of sales behaviors over the course of time based on over 140 different variables with a granularity of 25 observations per second. In addition, respondents (n=663) evaluated these presentations by means of a program analyzer with a granularity of 2 observations per second. Findings show that a substantial percentage of respondents form their impression about the sales representative within the first few seconds, whereas negative first impressions are formed faster than positive ones. The application of continuous measures (of nonverbal behaviors and customer responses) provides various advantages over existing means of measurement and yields important implications for marketers and future research.

4 Accounting for nonlinear, heterogeneous, and dynamic effects in store-level price response models *Philipp Aschersleben, Winfried J. Steiner*

It is well known that store-level brand sales may not only depend on contemporaneous variables like current own and competitive prices or other marketing activities, but also on past prices representing customer response to price changes. It has further been shown that accounting for lagged prices in a sales response model can increase expected brand profits over a static model that ignores price dynamics. On the other hand, non- or semiparametric regression models have been proposed in order to accommodate potential nonlinearities in price response, and related empirical findings indicate that price effects may show complex nonlinearities, which are difficult to capture with parametric models. Additionally, it is nowadays well established to incorporate store heterogeneity in sales response models, independent whether parametric or nonparametric modeling is used. We combine nonparametric price response modeling, heterogeneity and dynamic pricing. In particular, we model sales response flexibly using a Bayesian semiparametric approach and include the price of the previous period as well as further time-dependent effects. All nonlinear effects are modeled via P-splines, and embedding the semiparametric model into a hierarchical Bayesian framework further enables the estimation of store-specific (lagged) price response curves. In an empirical study, we demonstrate that our new model provides both more accurate sales forecasts and higher expected profits as compared to competing models that either ignore price dynamics or just include them in a parametric way. Optimal price policies for brands are determined by a discrete dynamic programming algorithm.

5 Was this review helpful to you? Determinants of helpfulness voting patterns in the context of online customer reviews

Filipe Sengo Furtado, Thomas Reutterer and Nadine Schröder

In recent years, the increase in user-generated content (UGC) has brought about a strong counterpart to information issued by manufacturers through marketing communication. With the rapidly increasing amount of customer reviews available online, 'helpfulness' features have been established to aid consumers in handling potential information overload. With this study we propose to deepen insights into what drives review helpfulness. While past research focuses on exclusively understanding what makes a review helpful and ignores the fact that some reviews receive more attention than others, we aim to disentangle these two dimensions by differentiating between what drives people to vote and what drives people to vote positively. Apart from well-known variables in the field of review helpfulness, such as review length, we focus on rating and text-related aspects in our research. This way, we are able to test the impact of different psychological and behavioral concepts (such as, e.g., consistency and conformity theory) on perceived helpfulness. We contribute to the existing literature by adopting a different modelling approach that enables us to separate two distinct effects that so far have been considered to be one. In doing so, we also identify a key determinant to the study of helpfulness.

6 Extracting Behavioral Intentions from Movie Trailer Comments: Which Video Components Matter to Consumers?

Jochen Hartmann, Jasper Schwenzow, Amos Schikowsky

In 2018, global box office revenues reached $42 billion (MPAA, 2018). Ample research has investigated how to forecast the commercial success of movies from pre-release predictors (e.g. Eliashberg et al (2000)), out of which movie trailers are the most important advertising tool. While few marketing scholars have explored the drivers of viral video ads (e.g., Nikolinakou and King (2018); Tellis et al (2019)), little knowledge exists about which video components of movie trailers matter most to consumers in forming behavioral intentions to watch a movie. Drawing on extant theories from cinematography and storytelling literature (e.g., Quesenberry and Coolsen (2019)), we propose a novel data analysis approach to establish a link between the video components of a movie trailer and consumer response. For this purpose, we pursue a multi-method approach. Specifically, we employ video mining and natural language processing techniques to analyze more than 1,000 movie trailers from YouTube of the highest-grossing English movies released in the years 2016-2018. To reveal behavioral intentions of consumers, we train a Random Forest (RF, Breiman (2001)) as a comment classifier to automatically detect "want-to-watch" expressions (e.g., "I can't wait to see this!!", "Finally..! Who else is going??"). RF is a versatile machine learning method, which can deal well with high-dimensional data (e.g., Hartmann et al (2019a); Hartmann et al (2019b); Wang et al (2018)). Classifying more than two million user comments, our analyses reveal a U-shaped effect between average trailer brightness and consumers' intention to watch a movie while brightness variance exhibits a negative association. Interestingly, we also find genre-specific interactions. Discussions about our findings with marketing managers from the movie industry suggest that our novel text-based success measure can complement existing success measures such as the number of comments and views to gain deeper knowledge about consumer response to movie trailers.

7 Confounding in Preference and Structural Heterogeneity *Narine Yegoryan, Daniel Guhl, Friederike Paetz, Daniel Klapper*

Consumer heterogeneity has been an important topic in choice modeling in marketing for many years. While the main focus has been on accounting for preference heterogeneity, only a few studies have recognized the importance of a specific type of structural heterogeneity, when consumers consider only a subset of attributes in a purchase decision (also referred to as attribute non-attendance). We use a latent class model with continuous parameter distributions in each class to account for both attribute non-attendance and preference heterogeneity. Restrictive cases of this model, ignoring either or both types of heterogeneity, enable us to investigate their possible confounding. Five empirical applications indicate that biases may arise in both cases either resulting in an overestimation of attribute non-attendance or biased estimation of preference heterogeneity. The results also suggest that the magnitude of the bias is application-specific and depends on the choice complexity and product category involvement.

8 A Mixed Logit model’s application: Personality traits as drivers for social preferences *Friederike Paetz*

Currently, social consumption constitutes a rapidly increasing trend that has great potential for companies. The characterization of social consumers is therefore highly relevant. To date, socio-demographic variables have been widely researched but turned out to be less appropriated to uniquely characterize social consumers. Psychographic variables are ascribed with the ability to overcome these problems, since recent studies maintain that consumers’ personal values and lifestyles are predictors of social consumption. However, personal values and lifestyle represent only two categories of psychological variables. Personality is another variable that is further known to be an antecedent of personal values and lifestyle. In this study, we focus on the characterization of social consumers based on their personalities. We conduct an empirical discrete choice experiment and investigate consumers’ personalities as a driver of consumer preferences for the fair trade (FT) label attribute. To operationalize consumers’ personalities, we use the popular five-factor approach. For the determination of consumer’s preferences, we estimate a Mixed Logit model that includes both unobserved preference heterogeneity and observed heterogeneity. Observed heterogeneity is captured by both consumer’s personality as well as socio-demographic variables. We find gender, academic degree and income as well as four personality traits as important drivers for consumers’ social preferences. We work out interaction effects between socio-demographic and personality variables and argue for the consideration of personality within the characterization of social consumers as the core sources for social preferences.

9 The Effect of Randomly Simulated Missing Value Patterns on PLS, ML and FIML Model Fit *Malek Simon Grimm, Ralf Wagner*

Missing values (MVs) are an omnipresent problem in quantitative and survey-based researches. Missing data hinder a researcher's ability to investigate a phenomenon of interest (McKnight et al, 2007) or lead substantial biases. Already a few missing values handled through a case exclusion (case-wise deletion, listwise deletion or complete case analysis) causes significant attribution of the total sample size: A data set with 500 observations and 10 variables with 10% MVs could, for example, reduce the effective sample size to 175 participants if a listwise deletion is applied (Cheema, 2014). A loss of data or information decreases statistical power and MVs can lead to biased results or estimates (Roth, 1994). Nowadays, several Missing Data Techniques (MDTs) are available. Unfortunately, the achievements in the statistical domain seem to have a negligible impact on research practices (McKnight et al, 2007). Building upon the studies of Parwoll and Wagner (2012) and Grimm and Wagner (2019), this research investigates the effects of MVs on measurement quality within structural equation modelling (SEM). Thereby, the measurement quality, accuracy and stability of the frequently used estimation methods partial least squares (PLS), maximum likelihood (ML) and full-information maximum likelihood (FIML) is comprehensively investigated. MV patterns within the range of 2.22% until 27.78% of MVs are implemented repeatedly into a data set for the European customer satisfaction index (ECSI). The simulation and comparison of repetitive random dropout mechanisms provides a robust understanding about the performance of PLS, ML and FIML.

10 Know your Limits: Requirements for the Application of MCMC Procedures for Pareto/NBD Distributed Data Sets *Lydia Simon*

The Pareto/NBD model is one of the best-known and most used models in customer base analysis. Still, practitioners are confronted with the question of which cohort size and length of calibration period are necessary in order to obtain reliable parameter estimates. In the past years, the usage of Monte Carlo Markov Chain (MCMC) algorithms has increased as these deliver a full posterior distribution rather than just a point estimate for the model parameters. Using MCMC additionally requires hyper parameters whose choice has barely been discussed in the literature yet. We, therefore, perform a broad simulation study on Pareto/NBD distributed data sets to derive minimal requirements for the model´s usage and to outline the choice and influence of different hyper parameters. The results show that the recovery of the purchase process already works well for cohort sizes of 1,000 customers and a calibration period of 52 weeks. Since we are in a non-contractual setting, the dropout process cannot be observed and is therefore much more difficult to estimate from the data. It requires a calibration period of at least two years and 5,000 customers. For all data sets, we generate MCMC estimates using different hyper priors as well as the uninformative Jeffreys´ prior. The goodness of fit measures tell us that that Jeffreys´ prior should be preferred to informative hyper distributions. This especially holds, when we have no preliminary information on our data set.

References

Breiman L (2001) Random Forests. Machine Learning 45(1):5–32.

Cheema JR (2014) A Review of Missing Data Handling Methods in Education Research. Review of Educational Research 84(4):487–508.

Eliashberg J, Jonker JJ, Sawhney MS, Wierenga B (2000) MOVIEMOD: An Implementable Decision-Support System for Prerelease Market Evaluation of Motion Pictures. Marketing Science 19(3):226–243.

Grimm MS, Wagner R (2019) The Impact of Missing Values on PLS, ML and FIML Model Fit. In press.
Hartmann J, Heitmann M, Schamp C, Netzer O (2019a) The Power of Brand Selfies in Consumer-Generated Brand Images. URL: www.brand-selfie.com. Working paper.
Hartmann J, Huppertz J, Schamp C, Heitmann M (2019b) Comparing Automated Text Classification Methods. International Journal of Research in Marketing 36(1):20–38.
Heckman J (1979) Sample Selection Bias as a Specification Error. Econometrica 47(1):153–161.
Kane K, Lo V, Zheng J (2014) Mining for the Truly Responsive Customers and Prospects Using True-lift Modeling: Comparison of New and Existing Methods. Journal of Marketing Analytics 2(4):218–238.
McKnight PE, McKnight KM, Sidani S, Figueredo AJ (2007) Missing data: A gentle introduction. Methodology in the social sciences. Guilford.
MPAA (2018) A Comprehensive Analysis and Survey of the Theatrical and Home Entertainment Market Environment (THEME) for 2018. URL: www.motionpictures.org/wp-content/uploads/2019/03/MPAA-THEME-Report-2018.pdf.
Nikolinakou A, King KW (2018) Viral Video Ads: Emotional Triggers and Social Media Virality. Psychology & Marketing 35(10):715–726.
Parwoll M, Wagner R (2012) The Impact of Missing Values on PLS Model Fitting, Springer, pp. 537–544.
Quesenberry KA, Coolsen MK (2019) Drama Goes Viral: Effects of Story Development on Shares and Views of Online Advertising Videos. Journal of Interactive Marketing 48:1–16.
Radcliffe N, Surry P (1999) Differential Response Analysis: Modeling True Response by Isolating the Effect of a Single Action. Credit Scoring and Credit Control IV.
Roth PL (1994) Missing data: A conceptual review for applied psychologists. Personnel Psychology 47(3):537–560.
Surry P, Radcliffe N (2011) Quality Measures for Uplift Models. Knowledge Discovery in Databases Conference 2011 (KDD-2011), August 21-24, 2011, San Diego, CA.
Tellis GJ, MacInnis DJ, Tirunillai S, Zhang Y (2019) What Drives Virality (Sharing) of Online Digital Content? The Critical Role of Information,

Emotion, and Brand Prominence. Journal of Marketing 83(4):1–20.
Wang Q, Nguyen TT, Huang JZ, Nguyen TT (2018) An Efficient Random Forests Algorithm for High Dimensional Data Classification. Advances in Data Analysis and Classification 12(4):953–972.